DATABASES FOR ACCELERATOR CONTROL – AN OPERATIONS VIEWPOINT

J. Poole, CERN, Geneva, Switzerland

Abstract

The data in an accelerator control system is not only important for the control processes but also serves as the primary source of information required for the understanding of the accelerator's behaviour. The management of this resource is therefore fundamental to accelerator performance and in recent years there have been technological developments which have made new approaches feasible. This paper reviews the impact of database and related technologies on accelerator operation and performance, their implementation and the implications of their use.

I. INTRODUCTION

Operating databases for accelerator control has a scope much wider than the control system and accelerator operation. The constraints and relationships from outside the control system are such that the databases have to be considered in the wider context of the whole project: this view will be reflected in this paper.

The term database is commonly used to describe many different things including simple file systems. However, here the term 'database system' will be used to describe something which satisfies the following:

- a structured collection of data, held in computer storage, and the software which is used to access the data in a selective way. This means that the user can access the relevant data without having to know how or where the data is stored.
- a system which will allow concurrent and consistent manipulation of the data by multiple users.

This definition essentially limits the field to commercial DataBase Management Systems (DBMS).

The size of control systems and the complexity of the data associated with them has become a driving force behind the use of databases. Database technology has reached its teen age and in the last few years its performance (and that of the platforms it runs on) has increased by orders of magnitude, making their use in control systems a viable option.

The use of database systems inside a control system can reduce the maintenance burden, improve understanding of the machine's behaviour and lead to improved performance. Furthermore, the most fundamental benefit which comes with the introduction of database systems is derived from the creation of a unique source for the data and this has far reaching benefits for an organisation. In this paper, the way in which database systems are implemented and used will be reviewed and their impact on accelerator performance will be explored. The requirement for a coherent project-wide solution is stressed.

II. DATA MANAGEMENT

Inside every control system there is data describing the settings and performance of the machine; being able to analyse and correlate it is fundamental to improving performance. A coherent solution to the data management problem is essential if many of the goals of accelerator operation are to be achieved.

Data is of little use if it cannot be accessed from a variety of places. At the same time, one requires tools to maintain and analyse it.

Today, commercial DBMS allow access from within control applications written in high level languages (C, Fortran etc.) and at the same time offer a variety of interfaces to commercial packages like data analysis tools and other databases.

Previously there were often many different data management systems at an accelerator site and this made control software difficult to develop and maintain because it involved diverse methods to access data. In a traditional system, a control program might have to address a power converter to find its current settings, an instrument to obtain a measurement and a database to obtain some calibrations: this could mean three completely different sets of routines within one small control program. The overheads in maintaining such software are very high: the individual systems all have to be maintained and when any one changes it implies changes to the control software as well.

Solutions based on databases where the access to the data is uniform can solve many of these problems. In addition, the peripheral problems like file management are greatly simplified. In some cases the advantage of lower maintenance overheads outweighs performance degradations arising from non-optimised systems. The database solution satisfies the requirement that a single data source is available to many users simultaneously in a heterogeneous environment and once this approach has been accepted, it leads to simpler systems which can be developed faster.

III. THE DATA UNIVERSE

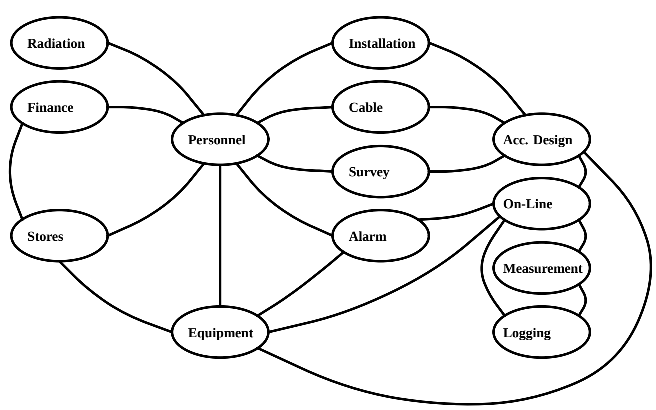

Figure. 1. A diagram representing some databases one would typically find at an accelerator laboratory and the links amongst them.

Although at many accelerator laboratories the database systems have evolved along with the various activities, the next generation of machines will feature project-wide, integrated database systems (see for example [1],[2]). In an accelerator laboratory there are gigabytes of data concerning personnel, finance, hardware and software and they are all inter-related; Figure 1 represents some of these inter-relationships, where the databases could possibly contain the following:

Radiation	Measurements, equipment (monitor) descriptions and calibrations
Finance	Purchasing, budgets, internal orders
Stores	Catalogue, suppliers
Personnel	Names, addresses, phone numbers, contractual details
Equipment	Inventory, procurement, testing, components, calibrations
Installation	Logistics, location of components, planning
Cables	Specifications, routing and connections
Survey	Position of elements and their evolution
Alarm	Descriptions, actions required, history, hierarchical information
Acc. Design	Accelerator components and assemblies, machine definition and historical evolution
On-line	Reference settings, actual settings, calibrations, data for applications
Measurement	Data from instrumentation updated at frequent intervals and kept for relatively short periods
Logging	Long term record of settings and measurements of the accelerator and its environment

The diagram illustrates the surprising links between conventional control system parameters and for example, personnel information. Such a link can be understood through the following example: a surveillance program detects an equipment fault; from the equipment data it can find the name of the person responsible; it can then deliver the phone numbers of this person to the accelerator operator who has to ensure that the fault is fixed.

The databases are separated according to the various functions of the laboratory and the different primary users of the information. This compartmentalisation is quite natural and necessary, particularly from the data maintenance aspect.

IV. SOCIOLOGY

The relationships among data at an accelerator facility cross many boundaries and it is a challenge to build systems in which these relationships can be exploited. The easiest way to achieve it, is for all databases to be implemented under the same DBMS or by having DBMS which can talk to each other. In a new project it is therefore important to reach an agreement on a common DBMS at an early stage so that developments in the different disciplines do not diverge.

Good design is a fundamental requirement for good performance from a DBMS and this in turn means that there must be a good initial information analysis. It has become generally accepted that flexible and reliable application software needs to be built by competent engineers who have been trained in appropriate methodologies. The same is true for databases but the technology is different and different expertise is required.

In order to achieve the goal of coherent database systems it is necessary to reach an agreement at project (or laboratory) management level on the policy. It is very easy for someone to believe that the data with which he is working is of no interest to anyone else and therefore he can use whatever system he has available to manage it; this inevitably leads to copies of files, inconsistencies and incompatibilities. Furthermore it is necessary to allocate adequate resources, not simply in terms of MIPS and DBMS licenses but also in personnel and their training.

Several categories of personnel are required to build and exploit databases: information analysts, database designers, software engineers and people who understand the particular system (personnel, hardware etc.). It is important that the people who understand the system are responsible for maintaining the data and therefore that they have some understanding of database technology. It is possible that one person could fulfill all roles for a particular system but it is unlikely that he could cover many more.

V. METHODOLOGY

A. Starting Point

Every new accelerator project starts with a computer simulation and this is an ideal point from which to start the data modeling and database implementation. The description of the new machine contains the names of the major components and the way in which they must be put together. This information is at the heart of nearly all other systems: the named components have some physical properties which form the basis for specifications for procurement; they have dimensions and positions which constrain the civil engineering; the components have to be connected to a control system and so on.

B. Analysis and Design

Many database instances have to be planned at an early stage in a project because they concern the hardware and contain information about procurement, construction, performance etc., others may come at a later date because they concern control data which are only needed by the application software. In both cases a detailed analysis of the information structures is essential before an effective database design can be established.

The analysis associated with the application software should be aimed at identifying the data which is necessary to define the states of the machine and software (evolution in time as well as snapshots). This leads to a better understanding of the machine behaviour [3]. The procedure requires formal methods like Structured Analysis and Structured Design (SASD) [4]. Powerful data modeling techniques are incorporated in SASD but they target the data involved in the software processes and therefore do not model the project-wide data. For information structures on the fringes of the control system, like the cable system, other techniques, such as the Nijssen Information Analysis Method (NIAM) [5], are more appropriate.

The fundamental point about this stage in the implementation is that analysis and design should be based on formal methods and supported by CASE tools. A good database design will lead to robust application software which can be developed rapidly whilst on the other hand, a simple error in design can easily cause a DBMS to stop.

C. Object Technology

Object oriented methods have been a feature of accelerator control systems for many years but the majority of projects have been, or are being, built using conventional database technology. There is an increasing interest in using object databases because they fit well with current software developments. This new type of database system is a direction which is being exploited by most vendors and it is possible that the technology will have reached maturity in time to be used in the next generation of accelerators.

VI. IMPLEMENTATION ISSUES

A. Replication

Making copies of the unique source of data is dangerous and goes against the fundamental principles of database implementation but it is often necessary for at least some of the data. The latest generation of DBMS provide tools for data replication and in some cases facilities to re-integrate modified copies. This recent development is another feature which has helped the proliferation of DBMS in accelerator laboratories because it allows more control over the copied data and automatic procedures for maintenance.

B. Limitations

A DBMS will never be as fast as a data server purpose-built for a given application, but in many applications the DBMS will be fast enough. The advantages of the DBMS have led to an increased range of applications in accelerator control with only the high performance software requiring special data management. Even in these cases, non-volatile data is often down-loaded from a DBMS because it is easier to maintain the data there, and at the same time processed data is passed upwards to be stored in a DBMS.

C. Database Servers

The following is a list of requirements for database servers of various types:

- 24 hour availability
- Close to 100% reliability
- High performance query server
- Very large tables
- Mostly write or mostly query
- Multi-user, interactive or single user, batch
- Reliable production or flexible development environment

It is highly probable that in any project one will want to have a system with high reliability and availability, for control for example, and at the same time a development system which will almost certainly crash fairly often. When one adds other systems with further conflicting requirements it is clear that a single server is unsuitable and one must therefore plan to have several platforms, each tuned to satisfy the particular requirements.

VII. DBMS PERFORMANCE

A. Speed

For some years the speed of transactions has been at a level where the use of DBMS inside a control application has been possible. Today on a typical control system platform, it is possible to run simple transactions (query, update) at rates ~ 100 Hz against quite large tables ($> 10^4$ rows). This means that the use of the database is transparent to the operator for the majority of operations.

B. Size

Existing DBMS can comfortably handle ~ 100 Gigabytes and the next versions are targeting Terabytes. This development has enabled huge amounts of data to be stored and manipulated with ease, leading to interesting new possibilities for analysis and correlation.

The construction of such very large databases requires some special techniques [6] like filling all rows in a table at creation time (often $> 10^6$ rows) and only updating non-indexed rows. An interesting feature of correlation in very large databases where the rows in different systems do not have one-to-one relationships was reported in [6]: it is much more efficient to extract the relevant data and do the correlation in a separate program. Another difficulty with very large databases concerns backup: one has to copy a complete table even if only one row in several million has been updated. Commercial DBMS vendors see very large databases as an important market for the future and are therefore improving in these areas.

C. Reliability

In the 1980's one would not have considered using a DBMS in a mission-critical role in a control system because they were too slow and because they were not reliable. This situation has certainly changed: the whole of CERN's accelerator complex relies on RDBMS for operation. The PS and LEP share a database server for on-line control and the Z physics program in LEP relies on data being available in the logging database.

VIII. DISADVANTAGES

In a control system there are usually a number of different platforms and a DBMS may not run on all of them. This is a more serious problem in the larger machines where the number of platforms concerned tends to be larger. The result is that software running on the non-standard platforms cannot access the data directly and different methods (e.g. remote procedure calls or local files) have to be employed.

Unfortunately there is an overhead associated with the use of DBMS: a support infrastructure is required because the database kernel requires tuning, the disk space managing, backups have to be done and the developers need advice. Also, as with all commercial software, one is obliged to install the latest releases in order to maintain support from the manufacturer. When there are multiple platforms with new versions of their operating systems arriving from time to time, combined with evolving versions of the DBMS software, the system management is non trivial.

IX. IMPACT ON ACCELERATOR PERFORMANCE

A. More Efficient Software

With a well designed data model at the heart of the control system it is easy to build generic applications. In LEP, for example, a single application is used to make trims to 90% of the machine parameters. Another application controls the status of almost all of the equipment. This obviously reduces the resources required to build and maintain the applications and provides a simple interface for the operator.

The well designed data model also leads to simple solutions for many problems e.g. a particular set of orbit correctors will often perform more than one task, like closed orbit correction and a local orbit bump. Through an understanding of the processes it is possible to identify (and label) the data required to define the two separate functions and their states, and build structures to store them. Once these two aspects are separately identifiable it is possible to build a save-and-restore utility which will allow selective restore of the bump and/or orbit. A good database implementation also makes it possible to perform many complex actions and to easily accommodate new operational requirements throughout the accelerator's life cycle.

B. Settings Management

Every accelerator needs to store and re-create previous runs but this can rarely be achieved by re-loading a file containing a particular set of values. In a collider for example, the performance depends on the whole history of operations from the starting point before the first beam was injected through accumulation, acceleration, β-squeezing and so on. Being able to reproduce a previous run requires the re-construction of all the actions; improving it, requires one to be able to modify a subset. Partitioning the data and being able to identify the various components by function, state and when they were used, allows selective retrieval for re-loading in the machine – no more thumbing through the log book to find hand written values or file names.

C. Understanding the Physics

We can use the data to improve our understanding of the physics of the accelerator by looking at what we have asked the machine to do and comparing it with beam measurements. A simple example of this concerns the magnetic model: we set a given tune value which is converted to magnet strengths through the model and these in turn are converted to currents for the power converters through the magnet calibrations. Analysis of this process leads to refinement of the model and therefore better performance.

Correlation can also lead to a better understanding of the physics. A very good example of this was in LEP when there were unexplained variations in the beam energy: it was possible to demonstrate that the magnetic fields were stable but that there was a correlation with temperature variations in the tunnel.

D. Building on the Database

Once a database has been set up it becomes possible to build applications which use the data and to feed the data into various other packages for processing. The immediate spin off from a database containing up-to-date measurements from the instrumentation is the ability to make arbitrary displays of beam parameters for the operators. Furthermore, combining this and the control database in an application allows one to close feedback loops.

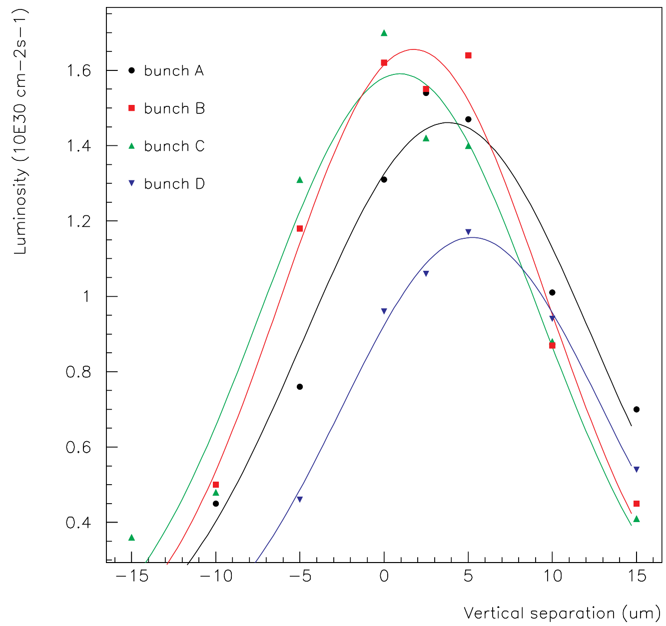

Figure. 2. Plot of a Vertical Luminosity Scan of Bunch Trains derived from the LEP Measurement Database

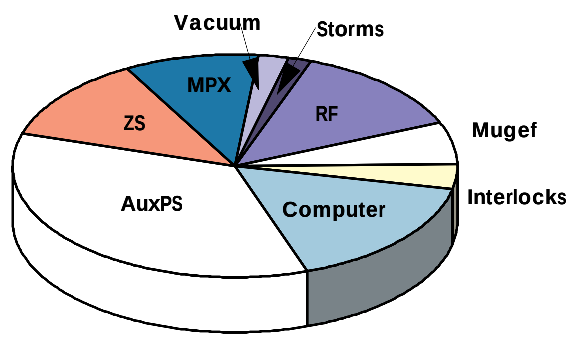

Figure. 3. Analysis of faults in the SPS which caused loss of physics in LEP. The data was recorded in the SPS Database, selected with SQL and plotted using Excel

Figure 2 shows a typical plot of some data taken from the LEP Measurement Database [7]. The advantage of having a single repository for the measurement data is that the programmer does not have to know anything about the hardware or how to retrieve the data from it; he only has to access the database. In fact, a generic plotting facility based on the measurement database has been written, and the operator only has to specify what to plot.

Figure 3 is an example of a very different use of the data stored in the control system. It shows the result of an analysis of the fault history recorded in the SPS statistics database and demonstrates the usefulness to managers of the general availability of the data.

At LEP there have been many performance improvements which can be directly attributed to the database systems. Indeed in the Proceedings of the LEP Performance Workshops [8] one can see that the percentage of graphs and tables derived from database data has increased from $<$ 5% in 1993 to $>$ 50% in 1995. People are able to access the data from many platforms and use their own favourite analysis package to process the data and this encourages them to study the data.

X. CONCLUSIONS

The commercial world has long regarded its information structures as corporate resources; in accelerator facilities this attitude is becoming more widely accepted. This implies a more global approach to information services and when properly implemented, leads to greater efficiency throughout the organisation.

It has been the improved performance of commercial RDBMS which has led to their proliferation in control systems. The speed and capability to handle large volumes of data is now adequate for many accelerator applications.

A coherent set of databases can be used for design, operation, maintenance and managerial applications. This approach leads to greater efficiency and improved accelerator performance whilst satisfying a wide variety of users but requires project-wide support and substantial resources.

Applying database technology in the way described above improves the understanding of the physics of an accelerator. At the same time managers can access performance data and analyse faults without having to search (or have someone else search) for data. Finally, the day-to-day operation of a machine becomes more reproducible and the complex functionality of an accelerator is better controlled.

XI. ACKNOWLEDGEMENT

I have drawn extensively on the experience of the people at CERN concerned with building the control databases for LEP and the SPS. The on-line control database was designed and built by P. Collier and M. Lamont; the logging and measurement databases by R. Billen, F. Bordry and E. Hatziangeli; the alarm database by M. Tyrrell and the vacuum by P. Strubin. I would also like to thank my colleagues at CERN and in the USA for their comments and ideas.

References

[1] A. Chan et al., ‘The PEP-II Project-Wide Database’, this conference.

[2] C.G. Trahern et al., ‘Relational Databases for RHIC Design and Control’, Proceedings of the Fourth EPAC, London, 1994, p. 1806.

[3] R. Bailey et al., ‘Development of the LEP High Level Control System using Oracle as an On-line Database’, Proceedings of ICALEPCS, Berlin, Nucl. Instr. and Methods in Physics Research A 352 (1994), p. 430.

[4] T. DeMarco,‘Strucured Analysis and System Specification’, Yourdon Press, New York, 1982.

[5] G. Verheijen and J. Van Bekkum, ‘NIAM: An Information Analysis Method’, 1982.

[6] R. Billen et al; ‘LEP accelerator logging system using an on-line database’, Proceedings of ICALEPCS, Berlin, Nucl. Instr. and Methods in Physics Research A 352 (1994), p. 296.

[7] M. Lamont,‘What has the Measurement System Given Us ?’, Proceedings of the Fifth Workshop on LEP Performance, CERN SL/95-08 (DI), February 1995, p. 23.

[8] Proceedings of the Workshops on LEP Performance, CERN SL/93-19 (DI), CERN SL/94-06 (DI) and CERN SL/95-08 (DI).